remote sensing

MDPI

Article

Deep Learning-Based Approximated Observation Sparse SAR Imaging via Complex-Valued Convolutional Neural Network

Zhongyuan Ji [1,2,3], Lingyu Li [1,3] and Hui Bi [1,3,*]

[1] College of Electronic Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China; jizhongyuan@nuaa.edu.cn (Z.J.); lly_815@nuaa.edu.cn (L.L.)
[2] College of Criminal Justice, Shandong University of Political Science and Law, Jinan 250014, China
[3] The Key Laboratory of Radar Imaging and Microwave Photonics, Ministry of Eduction, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
* Correspondence: bihui@nuaa.edu.cn

Abstract: Sparse synthetic aperture radar (SAR) imaging has demonstrated excellent potential in image quality improvement and data compression. However, conventional observation matrix-based methods suffer from high computational overhead, which is hard to use for real data processing. The approximated observation sparse SAR imaging method relieves the computation pressure, but it needs to manually set the parameters to solve the optimization problem. Thus, several deep learning (DL) SAR imaging methods have been used for scene recovery, but many of them employ dual-path networks. To better leverage the complex-valued characteristics of echo data, in this paper, we present a novel complex-valued convolutional neural network (CNN)-based approximated observation sparse SAR imaging method, which is a single-path DL network. Firstly, we present the approximated observation-based model via the chirp-scaling algorithm (CSA). Next, we map the process of the iterative soft thresholding (IST) algorithm into the deep network form, and design the symmetric complex-valued CNN block to achieve the sparse recovery of large-scale scenes. In comparison to matched filtering (MF), the approximated observation sparse imaging method, and the existing DL SAR imaging methods, our complex-valued network model shows excellent performance in image quality improvement especially when the used data are down-sampled.

Keywords: sparse synthetic aperture radar (SAR); sparse SAR imaging; iterative soft thresholding (IST); complex-valued convolutional neural network; deep learning

check for updates

Citation: Ji, Z.; Li, L.; Bi, H. Deep Learning-Based Approximated Observation Sparse SAR Imaging via Complex-Valued Convolutional Neural Network. *Remote Sens.* **2024**, *16*, 3850. https://doi.org/10.3390/rs16203850

Academic Editor: Massimiliano Pieraccini

Received: 20 August 2024
Revised: 6 October 2024
Accepted: 15 October 2024
Published: 16 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Recently, sparse synthetic aperture radar (SAR) imaging methods have achieved remarkable progress in image quality improvement, unambiguous reconstruction, and radar system simplification [1–3]. In 2007, Baraniuk and Steeghs stated that compressive sensing (CS) can replace matched filtering (MF) and reduce the required sampling rate at the receiver [4]. In 2010, Patel and colleagues introduced a new SAR imaging technique in the spotlight mode, which can provide a high-resolution reconstruction of the targets using significantly reduced data [5]. Then, Jiang et al. stated an innovative sparse imaging approach for directly processing the spaceborne SAR echo [6]. However, sparse SAR imaging algorithms based on exact observation matrix increase the computational complexity, which limits its application in the fast reconstruction of broad-scale scenes. To address this issue, Fang et al. formulated a new approximated observation-based SAR imaging method, which saves the computational costs in both time and memory [7]. In 2019, Bi et al. presented a frequency modulation continuous wave (FMCW) SAR sparse imaging method deducted from the wavenumber domain algorithm (WDA) [8]. This method compensates for the motion error in practical airborne SAR data, thereby achieving large-scale sparse imaging if the radar velocity is steady. Furthermore, Bi and co-authors introduced a novel real-time sparse SAR imaging method that reduces the sparse recovery time to a level

comparable to MF [9]. With the advancement in sparse SAR imaging technology, CS theory has been applied to the fields of three-dimensional (3-D) SAR [10,11] and inverse SAR (ISAR) imaging [12].

Nevertheless, CS-based SAR imaging algorithms still suffer from some issues, such as time-consuming iterative operations and optimal parameter selection, which hinder their further applications. Deep learning (DL) excels in feature learning and representation, providing a novel solution for addressing the challenges in sparse imaging. In 2017, Chierchia and co-authors applied convolutional neural networks (CNNs) to SAR image despeckling for the first time [13]. In the same year, Mousavi and Baraniuk developed a network called DeepInverse, which innovatively applied deep convolutional networks to sparse signal reconstruction [14]. In 2019, Gao and co-authors developed a radar imaging network employing complex-valued CNN, which was utilized to increase the imaging performance [15]. In 2020, Rittenbach and Walters presented an integrated SAR processing pipeline, which can map echo data to SAR images directly [16]. The deep neural network, RDAnet, performs SAR imaging and SAR image processing, ultimately achieving imaging performance comparable to the range Doppler algorithm (RDA). In 2021, Pu derived an auto-encoder model-based deep SAR imaging algorithm, and introduced a motion compensation scheme to eliminate errors [17]. In 2023, Meraoumia et al. proposed a self-supervised training strategy for single-look complex (SLC) images, named MERLIN, aimed at improving multitemporal filtering [18]. Lin et al. proposed a dynamic residual-in-residual scaling network for polarimetric SAR (PolSAR) image despeckling. Compared with traditional methods, this network is more computationally efficient and better preserves the image details [19].

However, the above methods treat the network as a "black box", which can bypass the optimal hyperparameter selection but lacks interpretable characteristics. Zhang and Ghanem developed ISTA-Net, a structured deep network for the CS reconstruction of SAR images, which converts the iterative soft thresholding (IST) algorithm into the deep network form [20]. On this basis, Yonel et al. presented a DL framework for passive SAR image reconstruction, which outperforms traditional sparse algorithms in both computational cost and recovered image quality [21]. In 2022, Zhang et al. proposed a proximal mapping model, named sparse representation-based ISTA-Net (SR-ISTA-Net), and applied it to the recovery of nonsparse scenes [22]. The SR-ISTA-Net is based on the one-dimensional (1-D) exact observation matrix, in which echoes and scene backscattering are expressed as vectors, making the storage and computing burden excessive. To overcome these shortcomings, Kang and colleagues introduced an innovative approximated observation-based SAR imaging net, which employs the range Doppler algorithm to construct the echo simulation operator, and achieves high-quality recovery of the surveillance area based on the DL imaging model [23]. Similarly, the deep network form of alternating direction method of multipliers (ADMM) has been implemented in SAR imaging. In 2021, Wei et al. casted multicomponent ADMM (MC-ADMM) into the deep network form, i.e., the parametric super-resolution imaging network (PSRI-Net), which obtains high-quality SAR images through end-to-end learning [24]. In 2022, Li et al. established a target-oriented SAR imaging model by unfolding the ADMM-based solution process, which improves the signal-to-clutter ratio (SCR) of the target [25]. Since the CNN module is not used for sparse representation, the above methods are not competent for nonsparse scene reconstruction. Then, Zhang et al. formulated a DL-based approximated observation SAR imaging method via the chirp-scaling algorithm (CSA) [26]. It has a dual-path CNN block in the nonlinearity module to enhance the sparse representation capability. In addition, sparse SAR imaging techniques based on deep unfolded network (DUN) have been widely used in fields such as SAR tomography (TomoSAR) [27,28] and ISAR [29], due to their enhanced reconstruction capabilities and efficiency.

Although the above-mentioned DL-based SAR imaging methods address the limitations of sparse data processing, they all employ path networks, which inevitably increases the complexity of the imaging model. This paper introduces a new complex-valued CNN-

based two-dimensional (2-D) sparse SAR imaging method. Our complex-valued network firstly constructs an approximated observation-based imaging model using the CSA. Then, the IST algorithm is utilized to address the L_1-norm regularization problem, and the problem solving procedure is mapped into a deep network form. Finally, a single-path complex-valued CNN block is constructed to increase the imaging performance. Compared with conventional approximated observation-based sparse imaging algorithms and existing DL SAR imaging networks, our imaging method shows superiority and robustness in scene reconstruction from both full- and down-sampled data.

The rest of this paper is structured as follows. Section 2 provides a brief introduction to the exact observation and approximated observation-based sparse SAR imaging models. In Section 3, our imaging network is introduced in detail, from the imaging model, the network structure, to the loss function. The experimental results based on the surface target and real scenes are provided in Section 4. Finally, conclusions are drawn in Section 5.

2. Sparse SAR Imaging

2.1. Sparse SAR Imaging-Based 1-D Observation Matrix

In this study, we adopt the airborne SAR imaging geometry depicted in Figure 1. It is assumed that the radar beam footprint has N_a pixels in azimuth and N_r pixels in range, $\mathbf{X} \in \mathbb{C}^{N_r \times N_a}$ denotes the 2-D backscattering coefficient matrix, and $\mathbf{Y} \in \mathbb{C}^{N_\tau \times N_\eta}$ denotes the 2-D echo matrix. Next, $\mathrm{x} = \mathrm{vec}(\mathbf{X}) \in \mathbb{C}^{N \times 1}$ with $N = N_r \times N_a$ represents the backscattering coefficient vector. Similarly, $\mathbf{y} = \mathrm{vec}(\mathbf{Y}) \in \mathbb{C}^{M \times 1}$ with $M = N_\tau \times N_\eta$. Considering the down-sampling processing of echo data, the exact observation-based sparse SAR reconstruction model can be formulated as [5]

$$\mathbf{y} \;=\; \mathbf{\Xi\Theta x} \;+\; \mathbf{n_0} \;=\; \mathbf{\Phi x} \;+\; \mathbf{n_0} \tag{1}$$

where $\mathbf{\Xi}$ is the sampling matrix, $\mathbf{\Theta}$ is the observation matrix, and $\mathbf{\Phi}$ is the system measurement matrix. When the considered scene is sparse enough and $\mathbf{\Phi}$ satisfies the restricted isometry property (RIP) condition, we achieve sparse imaging by solving [9]

$$\hat{\mathbf{x}} \;=\; \min_{\mathbf{x}} \left\{ \|\mathbf{y} - \mathbf{\Phi x}\|_2^2 \;+\; \lambda \|\mathbf{x}\|_1 \right\} \tag{2}$$

where λ is the regularization parameter, and $\hat{\mathbf{x}}$ is the reconstructed 1-D backscattering coefficient.

2.2. Sparse SAR Imaging-Based 2-D Observation matrix

The approximated observation-based sparse SAR imaging model can be expressed as [1,7,9]

$$\mathbf{Y} \;=\; \mathbf{\Xi}_a \circ \mathcal{G}(\mathbf{X}) \circ \mathbf{\Xi}_r \;+\; \mathbf{N} \tag{3}$$

where $\mathbf{N}$ is the noise matrix, $\circ$ is the Hadamard product operator, and $\mathbf{\Xi}_a$ and $\mathbf{\Xi}_r$ are the azimuth and range down-sampling matrices, respectively. Let $\mathcal{R}$ denote the imaging process of the MF SAR imaging algorithm, like CSA used in this paper, and $\mathcal{G}$ represent the inverse process of $\mathcal{R}$, named the echo simulation operator. The operators $\mathcal{R}$ and $\mathcal{G}$ can be defined as [30]

$$\mathcal{R}(\mathbf{Y}) \;=\; \mathbf{F}_a^{-1}\Big(\mathbf{H}_{ac} \circ \Big(\mathbf{H}_{rc} \circ (\mathbf{H}_{sc} \circ (\mathbf{F}_a\mathbf{Y})\mathbf{F}_r)\mathbf{F}_r^{-1}\Big)\Big) \tag{4}$$

$$\mathcal{G}(\mathbf{X}) \;=\; \mathbf{F}_a^{-1}\Big(\mathbf{H}_{sc}^{*} \circ \Big(\mathbf{H}_{rc}^{*} \circ (\mathbf{H}_{ac}^{*} \circ (\mathbf{F}_a\mathbf{X})\mathbf{F}_r)\mathbf{F}_r^{-1}\Big)\Big) \tag{5}$$

where $\mathbf{F}_a$ and $\mathbf{F}_r$ are the azimuth and range fast Fourier transform (FFT) matrices, respectively, $\mathbf{F}_a^{-1}$ and $\mathbf{F}_r^{-1}$ are the azimuth and range inverse FFT matrices, $(\cdot)^*$ is the conjugate transpose operator, $\mathbf{H}_{sc}$ is the chirp scaling operator, $\mathbf{H}_{rc}$ is the phase matrix used for the range compression and consistent range correction, and $\mathbf{H}_{ac}$ is the phase matrix used for

the azimuth compression and residual phase compensation. Then, the 2-D scene of interest is recovered by addressing the L_1-norm regularization issue [9]

$$\hat{\mathbf{X}} = \min_{\mathbf{X}} \left\{ \|\mathbf{Y} - \Xi_a \circ \mathcal{G}(\mathbf{X}) \circ \Xi_r\|_{\mathrm{F}}^2 + \lambda \|\mathbf{X}\|_1 \right\} \tag{6}$$

where $\hat{\mathbf{X}}$ denotes the scene of interest for reconstruction.

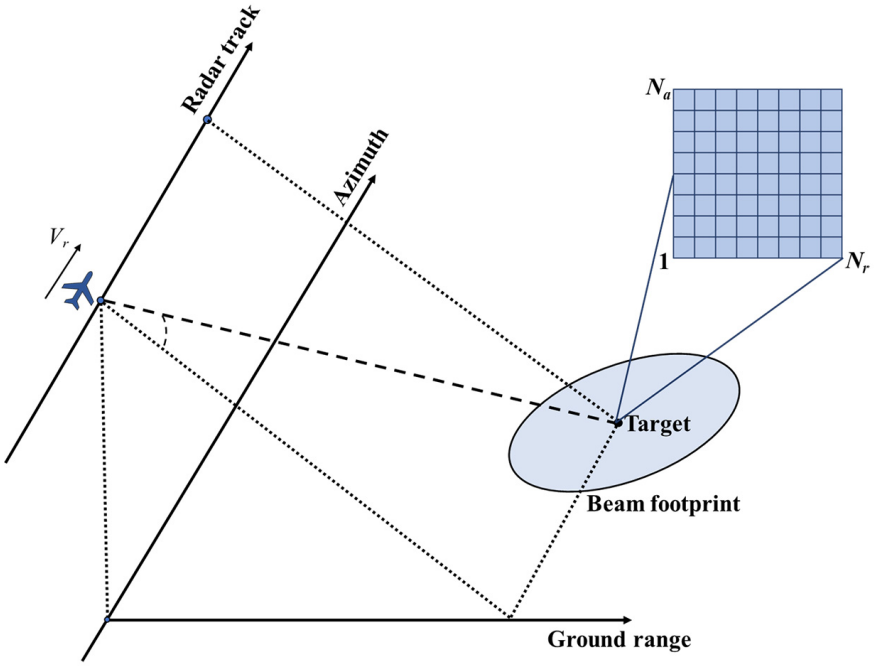


Figure 1. SAR imaging geometry.

3. Complex-Valued Network-Based Approximated Observation Sparse SAR Imaging

3.1. Complex-Valued Network Structure

There are several algorithms used to address the L_1-norm optimization issue, such as IST, ADMM [31], complex approximated message passing (CAMP) [32], and BiIST [33]. In this study, IST is selected to tackle the optimization issue presented in (6), whose iteration is among the following update steps

$$\mathbf{R}^{(k)} = \mathbf{X}^{(k-1)} + \rho \mathcal{R}\left(\mathbf{Y} - \Xi_a \circ \mathcal{G}\left(\mathbf{X}^{(k-1)}\right) \circ \Xi_r\right) \tag{7}$$

$$\begin{aligned} \mathbf{X}^{(k)} &= \mathrm{soft}\left(\mathbf{R}^{(k)}; \mathrm{T}\right) \\ &= \frac{\mathbf{R}^{(k)}}{\left|\mathbf{R}^{(k)}\right|} \max\left[\left(\left|\mathbf{R}^{(k)}\right| - \mathbf{T}\right); 0\right] \end{aligned} \tag{8}$$

where $\mathbf{R}^{(k)}$ denotes the 2-D linear reconstruction result in the *k-th* iteration, k denotes the IST iteration index, ρ represents the step size, and T denotes the threshold. However, IST requires multiple iterations to converge, with all parameters including ρ and T being pre-defined, which makes it difficult to adjust the prior information. Therefore, we map the update steps of IST to a deep network structure consisting of a fixed number of stages, each corresponding to an iteration in IST. In addition, considering that real scenes and their transformation domain results may not exhibit sparsity, we exploit CNNs to obtain a sparse representation of the data. In recent years, complex-valued neural networks have shown their effectiveness in image classification, speech spectrum prediction, and other applications [34]. Both the input echo and the output image for SAR imaging are 2-D complex matrices. Therefore, in order to effectively extract the complex-valued features, we choose a complex-valued neural network, which is represented by $\mathcal{F}(\cdot)$, whose parameters

are learnable. Then, the scene imaging of the complex-valued network can be divided into linear module R and nonlinear module N, which is written as

$$\begin{aligned} \mathrm{R}: \mathbf{R}^{(k)} &= \hat{\mathbf{X}}^{(k-1)} + \rho^{(k)} \mathcal{R}\left(\mathbf{Y} - \Xi_a \circ \mathcal{G}\left(\hat{\mathbf{X}}^{(k-1)}\right) \circ \Xi_r\right) \\ \mathrm{N}: \hat{\mathbf{X}}^{(k)} &= \tilde{\mathcal{F}}\left(\mathrm{soft}\left(\mathcal{F}\left(\mathbf{R}^{(k)}\right); \mathrm{T}^{(k)}\right)\right) \end{aligned} \tag{9}$$

where $\hat{\mathbf{X}}^{(k)}$ is the reconstructed 2-D scene, $\mathrm{T}^{(k)}$ and $\rho^{(k)}$ are the adaptive step size and the threshold in the *k-th* phase, respectively, and $\tilde{\mathcal{F}}(\cdot)$ is structured to be symmetric to $\mathcal{F}(\cdot)$. Specifically, $\mathcal{F}(\cdot)$ can be formulated as

$$\mathcal{F}\left(\mathbf{R}^{(k)}\right) = \mathcal{C}_2^{(k)}\left\langle \mathbb{C}\mathrm{ReLU}\left\{\mathcal{B}^{(k)}\left[\mathcal{C}_1^{(k)}\left(\mathbf{R}^{(k)}\right)\right]\right\}\right\rangle \tag{10}$$

where $\mathcal{C}_1^{(k)}$ is the complex-valued convolution operation, which uses N_f filters (each of size 3×3 in the experiments of this paper). This operation aims to convert a single-channel input into N_f channel outputs, and $\mathcal{C}_2^{(k)}$ is another set of N_f filters. Assume that $\mathbf{Z}$ is a complex-valued convolution matrix, $\mathbf{U}$ is a complex-valued image, and let $\mathrm{u} = \mathrm{vec}(\mathbf{U})$. Thus, the complex-valued convolution operation can be expressed as [34]

$$\mathbf{Z} * \mathrm{u} = \begin{bmatrix} \mathrm{Re}(\mathbf{Z}) & -\mathrm{Im}(\mathbf{Z}) \\ \mathrm{Im}(\mathbf{Z}) & \mathrm{Re}(\mathbf{Z}) \end{bmatrix} * \begin{bmatrix} \mathrm{Re}(\mathrm{u}) \\ \mathrm{Im}(\mathrm{u}) \end{bmatrix} \tag{11}$$

Next, $\mathcal{B}^{(k)}$ is the complex-valued batch normalization (BN) operation [34], which is defined as

$$\mathcal{B}(x) = \nu x + \beta \tag{12}$$

with

$$\nu = \begin{pmatrix} \nu_{rr} & \nu_{ri} \\ \nu_{ri} & \nu_{ii} \end{pmatrix} \tag{13}$$

where x is the input, β is the bias parameter with two trainable elements (the real and imaginary parts), and ν is the scaling parameter with three trainable elements. Then, $\mathbb{C}\mathrm{ReLU}(\cdot)$ is the complex-valued activation function, which can be expressed as [34]

$$\mathbb{C}\mathrm{ReLU}(\cdot) = \mathrm{ReLU}(\mathrm{Re}(\cdot)) + i\mathrm{ReLU}(\mathrm{Im}(\cdot)) \tag{14}$$

Therefore, the structure of our complex-valued network-based approximated observation sparse SAR imaging model is depicted in Figure 2.

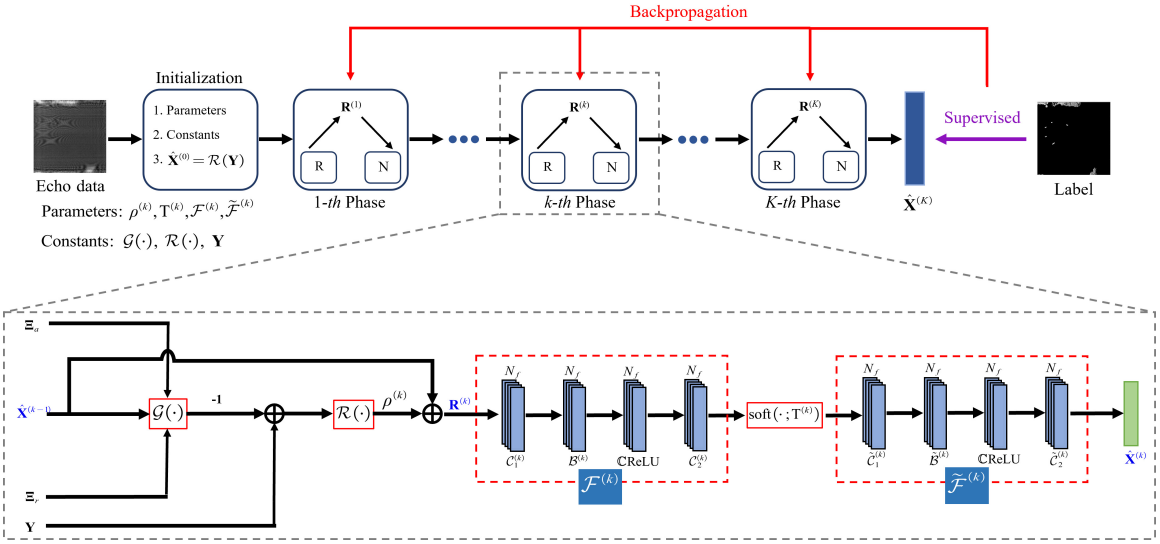


Figure 2. Network structure of our imaging method.

3.2. Loss Function Design

The training process for the complex-valued network involves updating the learnable parameters to minimize the loss function. In order to satisfy the symmetry constraints $\tilde{\mathcal{F}}(\cdot) \times \mathcal{F}(\cdot) = \mathbf{I}$, we adopt a loss function including the reconstruction error $\mathcal{L}_1$ and the symmetry constraint error $\mathcal{L}_2$ to design a novel total loss function, i.e.,

$$\mathcal{L}_{total} = \mathcal{L}_1 + \gamma \mathcal{L}_2 \tag{15}$$

with

$$\begin{cases} \mathcal{L}_1 = \frac{1}{2M} \sum_{m=1}^{M} \left\| \hat{\mathbf{X}}_m^{(K)} - \mathrm{D}_m \right\|_{\mathrm{F}}^2 \\ \mathcal{L}_2 = \frac{1}{2M} \sum_{m=1}^{M} \sum_{k=1}^{K} \left\| \tilde{\mathcal{F}}^{(k)} \left[\mathcal{F}^{(k)} \left(\mathbf{R}_m^{(k)} \right) \right] - \mathbf{R}_m^{(k)} \right\|_{\mathrm{F}}^2 \end{cases} \tag{16}$$

where $\mathbf{R}_m^{(k)}$ is the linear reconstruction result in the *k-th* phase corresponding to the *m-th* training sample; $\hat{\mathbf{X}}_m^{(K)}$ and D_m are the reconstructed result and label for the *m-th* training sample; $m = 1, 2, \cdots, M$; M is the total number of training samples; and γ is a weighting factor.

3.3. Complex-Valued Network Analysis

The detailed workflow of the constructed complex-valued network is illustrated in Figure 3, which takes 2-D echo data as input and outputs a 2-D SAR image. The primary innovation of this paper lies in the single-path network structure, which includes linear modules (denoted as R), nonlinear modules (denoted as N), and loss functions. Different from the dual-path linear module of CSA-Net and SR-CSA-Net, the proposed complex-valued network employs a single-path linear module to simplify the imaging model. The nonlinear module of the model includes complex-valued convolution, complex-valued batch normalization (BN), and complex-valued activation function, which treats the complex data as an entirety. $\theta = \left\{ \mathrm{T}^{(k)}, \rho^{(k)}, \mathcal{F}^{(k)}, \tilde{\mathcal{F}}^{(k)} \right\}_{k=1}^{K}$ is a learnable element set of the proposed method, where K is the total number of complex-valued network phases. Then, we optimize the learnable parameter set θ through supervised learning. It should be noted that trainable parameters across different phases are not shared. The number of learnable parameters in complex-valued CNN modules $\mathcal{F}^{(k)}$ and $\tilde{\mathcal{F}}^{(k)}$ is the same as $1 \times \omega_f \times \omega_f \times N_f + N_f \times \left(5 + \omega_f \times \omega_f\right) \times N_f$. Subsequently, the total number of learnable parameters in the designed complex-valued network can be written as

$$O_\theta = K \times \left\{ 2 + 2 \times \left[1 \times \omega_f \times \omega_f \times N_f + N_f \times \left(5 + \omega_f \times \omega_f\right) \times N_f \right] \right\}. \tag{17}$$

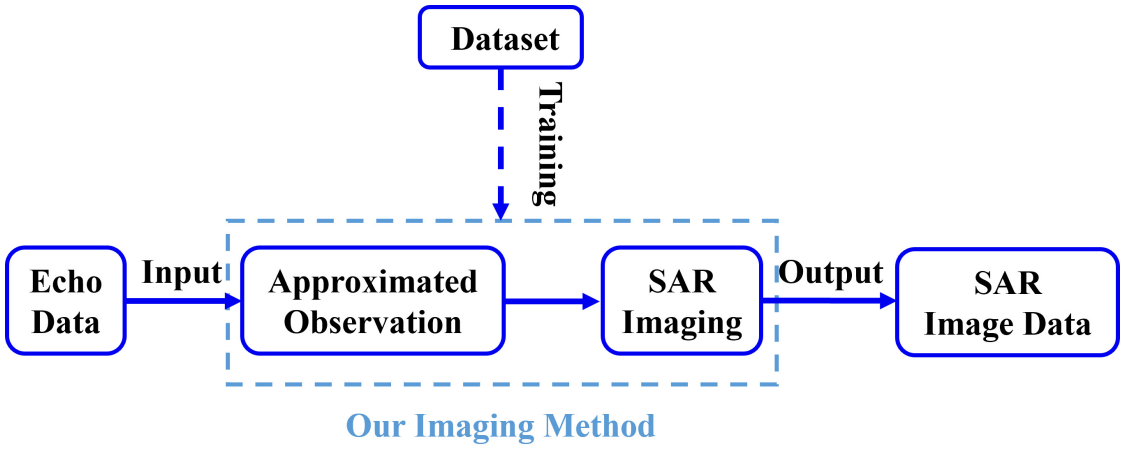


Figure 3. Complex-valued network workflow.

4. Experiments Based on Surface Target

The experiments using simulated data are conducted to verify the designed complex-valued network. Experimental parameters are shown in Table 1. All experiments are performed in the PyTorch framework with the Adam optimizer and are accelerated by an NVIDIA GeForce RTX 4090 GPU. The surface target simulated scene, which is the label of this part, is illustrated in Figure 4. During dataset generation, the full-sampled echo data are collected firstly. We then add some noise with a signal-to-noise ratio (SNR) randomly distributed between −10 dB and 35 dB to the collected echoes, and perform the down-sampling for the data to create the used simulated dataset. The down-sampling ratio (DSR) is randomly varied between 0.36 and 1. A total of 1600 simulated echoes with known imaging geometries and parameters are generated, of which 1280 are used for training and 320 for testing.



Figure 4. Surface target simulated scene. Horizontal axis is range direction.

Table 1. Simulated Parameters.

Parameter	Value
Carrier frequency	9.4 GHz
Pulse repetition frequency (PRF)	120 Hz
Effective radar velocity	150 m/s
Pulse duration	2.5 us
Bandwidth	100 MHz
Platform height	10^4 m
Number of phases (K)	7
Number of filters (N_f)	16
Learning rate	5×10^{-4}
Batch size	32

4.1. Anti-Noise Simulations

To explore the influence of SNR, full-sampled echoes with different SNRs are applied to CSA, an approximated observation-based sparse SAR imaging method (L_1-De) [9], CSA-Net [26], SR-CSA-Net [26], and our complex-valued network model. Figure 5 shows the recovered images of five different SAR imaging methods from full-sampled data under different SNRs. Due to the limited noise immunity of CSA, the target scene will be overwhelmed by noise at a low SNR. In contrast, L_1-De and CSA-Net have a certain anti-noise ability, but cannot maintain the backscattering coefficient of the target at low SNRs. However, SR-CSA-Net and our complex-valued network can reconstruct the backscattering coefficient of the target while suppressing noise. When SNR is −5 dB, the recovered image of the proposed method has a clearer outline than the SR-CSA-Net-based result (see Figure 5s,t).

In addition, the normalized mean square error (NMSE) and peak SNR (PSNR) are used to quantitatively assess the performance of different methods. The NMSE is formulated as

$$\mathrm{NMSE} = \frac{\|\hat{\mathrm{X}} - \mathrm{D}\|_{\mathrm{F}}^2}{\|\mathrm{D}\|_{\mathrm{F}}^2} \tag{18}$$

where $\hat{\mathrm{X}}$ is the recovered result, and D represents the label. Then, the PSNR can be expressed as

$$\mathrm{PSNR} = 20 \cdot \log_{10}\left(\frac{I_{max}}{\sqrt{\mathrm{MSE}}}\right) \tag{19}$$

with

$$\mathrm{MSE} = \frac{1}{N_a N_r} \sum_{n_a=1}^{N_a} \sum_{n_r=1}^{N_r} \left[\hat{\mathrm{X}}(n_a, n_r) - \mathrm{D}(n_a, n_r)\right]^2 \tag{20}$$

where MSE is the mean square error between the reconstructed result and the label, I_{max} is the maximum pixel value of the image, which is 255 in this paper. A smaller NMSE indicates that the reconstructed result is closer to the label, which reflects better imaging performance. The PSNR represents the similarity and distortion between the recovered image and the label.

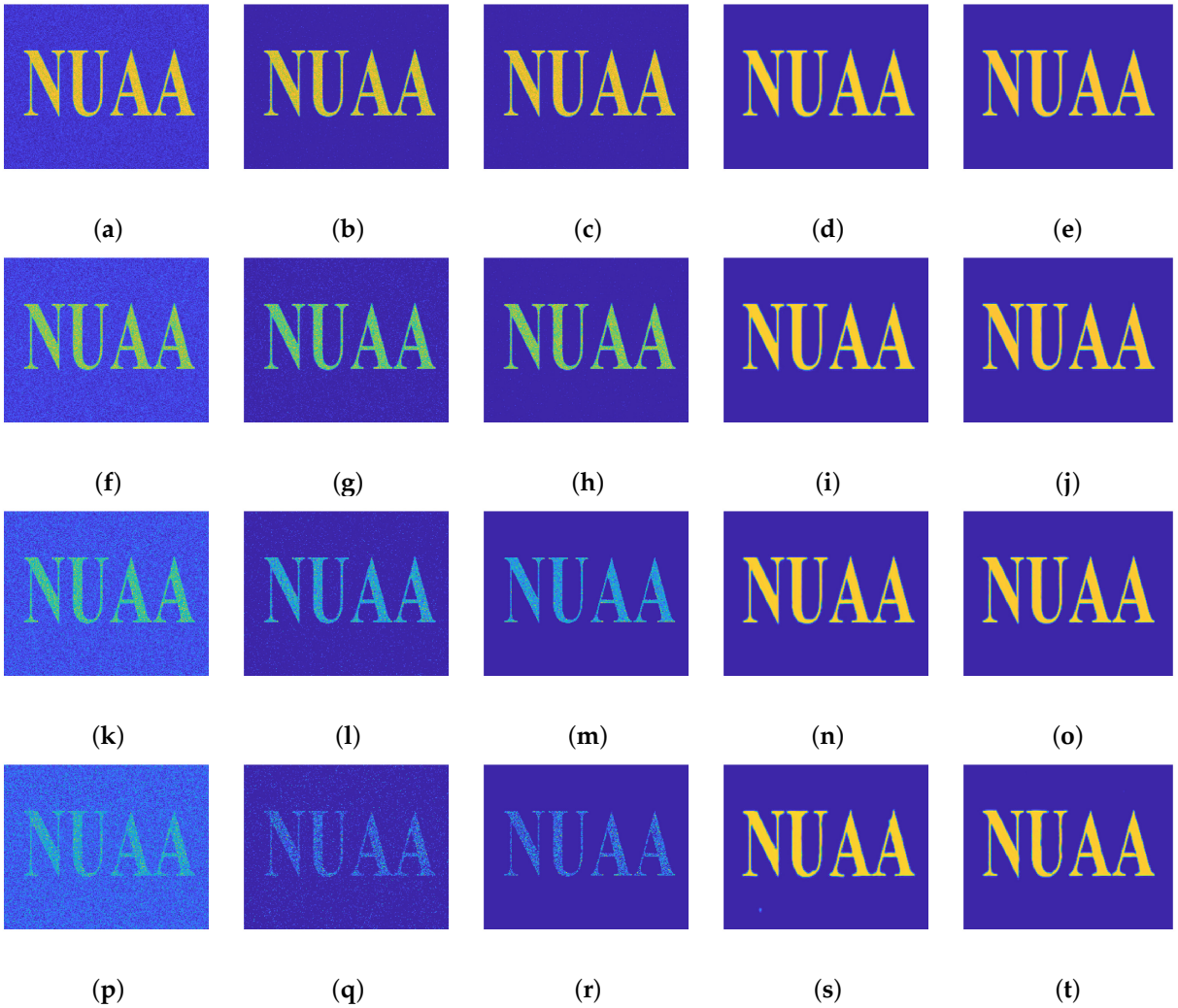


Figure 5. Recovered images of surface target from full-sampled echo by (from left to right column) (**a**,**f**,**k**,**p**) CSA, (**b**,**g**,**l**,**q**) L_1-De, (**c**,**h**,**m**,**r**) CSA-Net, (**d**,**i**,**n**,**s**) SR-CSA-Net, and (**e**,**j**,**o**,**t**) our complex-valued network, respectively. From top to bottom row, the SNRs are 10 dB, 5 dB, 0 dB, and −5 dB, respectively.

The average values of the quantitative indicators obtained by 100 Monte Carlo simulations are shown in Table 2. CSA, as a more traditional SAR imaging method, does not use optimization or DL techniques. Therefore, its NMSE and PSNR performance results are the worst among all methods. It is seen that compared with L_1-De, CSA-Net has lower NMSE and higher PSNR values in each case due to the optimal parameters obtained by training. Table 2 verifies that SR-CSA-Net and our complex-valued network significantly outperform the other methods. Our complex-valued network consistently shows the lowest NMSE values and the highest PSNR values, indicating its effectiveness in suppressing noise and preserving important image features, even under challenging conditions.

Table 2. Performance comparison with different SNRs.

	CSA		**L_1-De**		**CSA-Net**		**SR-CSA-Net**		**Complex-Valued Network**	
	NMSE	**PSNR (dB)**	**NMSE**	**PSNR (dB)**	**NMSE**	**PSNR (dB)**	**NMSE**	**PSNR (dB)**	**NMSE**	**PSNR (dB)**
SNR = 10 dB	0.2031	18.42	0.1500	19.73	0.1445	19.90	0.0042	35.28	0.0028	37.08
SNR = 5 dB	0.3193	16.45	0.2106	18.26	0.1861	18.79	0.0043	35.13	0.0036	35.93
SNR = 0 dB	0.5559	14.04	0.4078	15.39	0.3830	15.66	0.0064	33.43	0.0048	34.70
SNR = −5 dB	0.8776	11.44	0.7103	12.98	0.6730	13.21	0.0165	29.31	0.0128	30.43

4.2. Simulations Based on Down-Sampled Data

Then, we study the effect of different DSRs on the recovered image quality. To preserve generality, we add an additional 25 dB of white Gaussian noise to the collected echo. The reconstructed results of five SAR imaging methods under different DSRs are shown in Figure 6. It is seen that affected by the down-sampling, the CSA-based results have obvious energy dispersion and a defocusing phenomenon in both the azimuth and range directions. When DSR = 0.72 and DSR = 0.49, both L_1-De and CSA-Net are capable of reconstructing the simulated scene using the down-sampled echoes. However, they do not perform well at low DSRs, especially in the precise recovery of target scattering intensity. SR-CSA-Net can recover target scattering intensity but cannot eliminate the ambiguity phenomenon caused by the data down-sampling. For the proposed method, it achieves well-focused images in all cases.

The average values of the quantitative indicators of the recovered images in Figure 6 are shown in Table 3. The quantitative metrics of CSA, L_1-De, and CSA-Net are significantly declined by down-sampling. Therefore, for scenes with lower sparsity, the optimization technique struggles with accurate reconstruction from the incomplete data. Compared with CSA-Net, SR-CSA-Net further advances the performance of SAR imaging by incorporating CNN modules. Even at lower DSRs, SR-CSA-Net maintains relatively high reconstruction quality. The results show that although the imaging performance of our complex-valued network decreases with the decrease in DSR, it achieves the optimal NMSE and PSNR values in all test scenarios. Since the surface target simulated scene is fixed, the required computational time is approximately under different DSRs. It is seen that compared with L_1-De, the imaging time of three SAR imaging networks is closer to that of CSA. Among these three networks, CSA-Net has the shortest computation time attributed to the absence of sparse representation modules. Under down-sampled cases, our complex-valued network offers a compelling balance of accuracy and efficiency, making it suitable for the sparse imaging of broad-scale scenes.

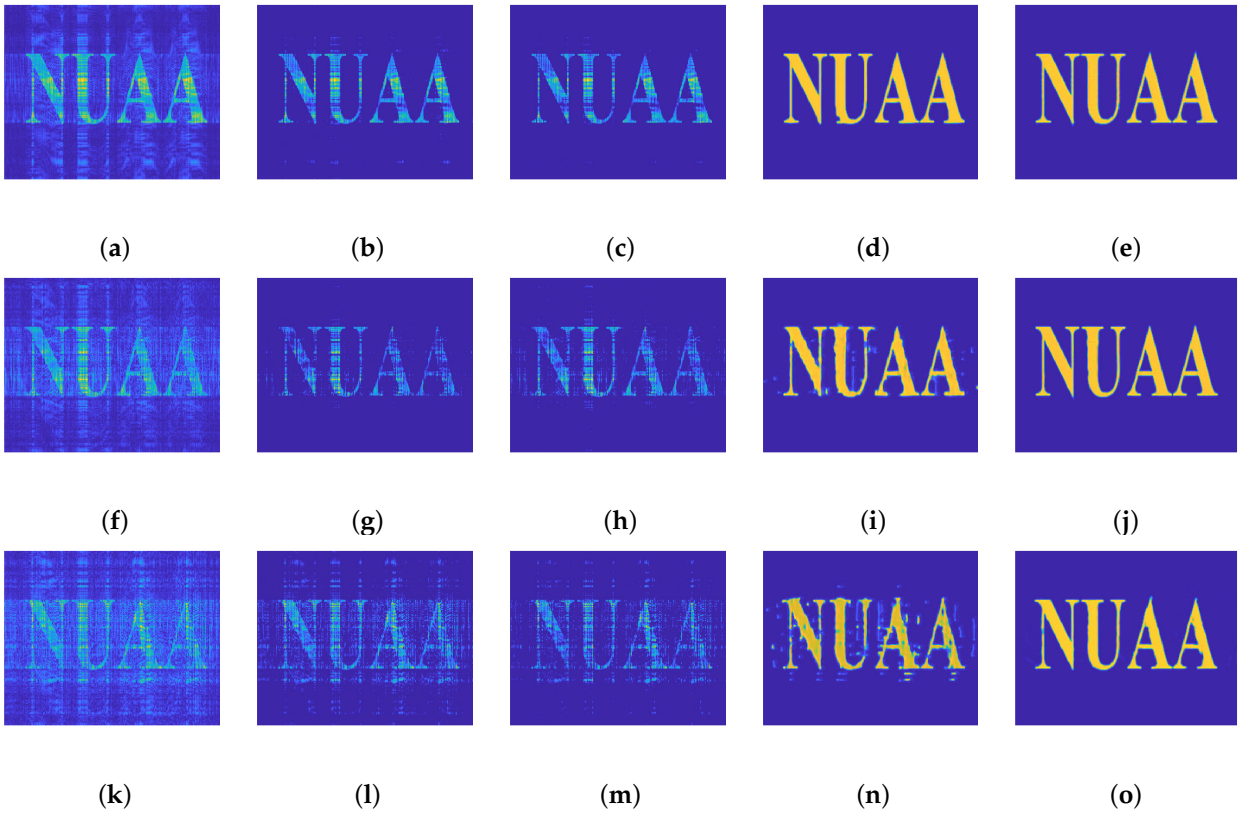


Figure 6. Recovered images of surface target from down-sampled echo by (**a**,**f**,**k**) CSA, (**b**,**g**,**l**) L_1-De, (**c**,**h**,**m**) CSA-Net, (**d**,**i**,**n**) SR-CSA-Net, and (**e**,**j**,**o**) our complex-valued network, respectively. From top to bottom row, the DSRs are 72%, 49%, and 36%, respectively.

Table 3. Performance comparison with different down-sampling ratios.

Method	DSR = 72%		DSR = 49%		DSR = 36%		Time (ms)
	NMSE	PSNR (dB)	NMSE	PSNR (dB)	NMSE	PSNR (dB)	
CSA	0.4585	14.95	0.4926	14.57	0.6209	13.56	3.5
L_1-De	0.4126	15.34	0.4763	14.71	0.5741	13.90	154
CSA-Net	0.3660	15.86	0.4699	14.77	0.5582	14.02	7.2
SR-CSA-Net	0.0130	30.35	0.0423	25.23	0.1278	20.43	13.6
The proposed method	0.0052	34.34	0.0076	32.69	0.0199	28.51	19.4

4.3. Comparative Experiments

To further show the superiority of our complex-valued network, the comparative experiments are presented in this part. Figure 7 illustrates the average PSNR curves at different phases and epochs when the DSR is 0.72. It is found that the average PSNR curves initially increase with the number of phases and then plateaus. Clearly, when K is set to 7, SR-CSA-Net improves the PSNR by nearly 15 dB over CSA-Net, and the proposed method achieves about 4 dB additional gain over SR-CSA-Net. Figure 7b indicates that SR-CSA-Net and the proposed method obtain higher PSNR as the number of training epochs increases. Moreover, the proposed method demonstrates superior performance upon training convergence. In summary, the symmetric neural network modules used in SR-CSA-Net and our method significantly enhance the performance of sparse SAR imaging. In addition, balancing network complexity with reconstruction performance, we set $K = 7$ and the default epoch to 150, which provide optimal convergence.

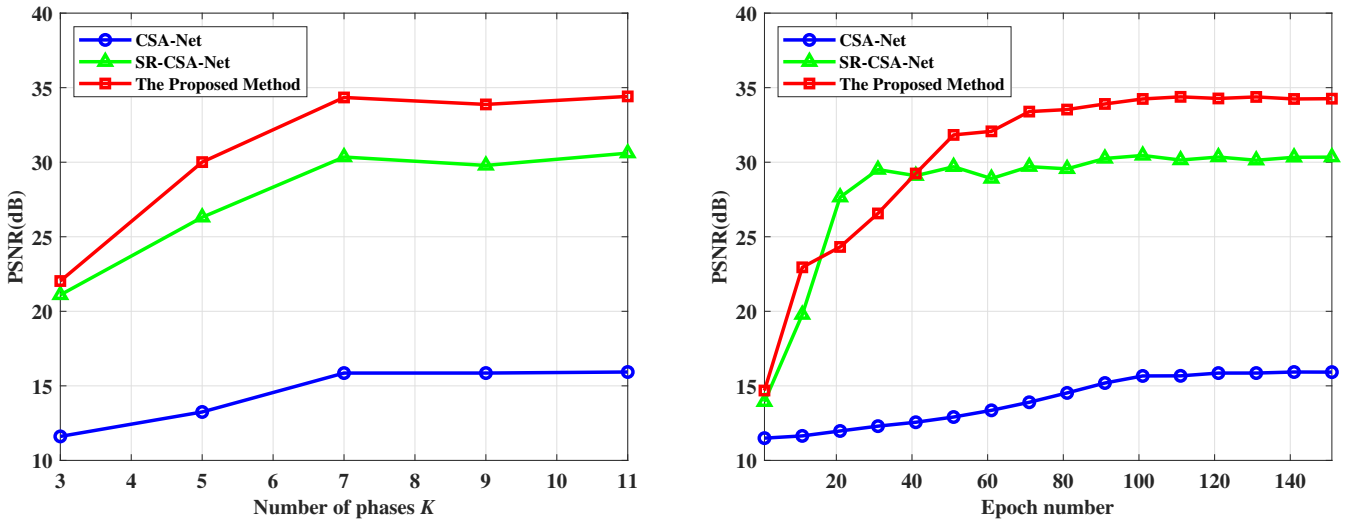


Figure 7. PSNR comparison of three SAR imaging networks with various numbers of (**a**) phases and (**b**) epochs.

5. Experiments Based on SSDD Dataset

In this section, we use the Open SAR Ship Detection Dataset (SSDD) to prove the feasibility of our complex-valued network in a real scene [35] . The SSDD contains 1160 images, 928 of which are used for training. Firstly, the scenes in SSDD are cut into the size of 256×256 for experiments. Next, to preserve the target's scattering characteristics while minimizing background noise and clutter, we apply the maximum connection domain algorithm to obtain a suitable mask for each SAR image. It is worth noting that the mask assigns the target value to 1 and the background value to 0, using the mask to multiply the original real scene to obtain the training label as illustrated in Figure 8. Lastly, the dataset is created in a similar way to the previous simulation, with the only difference being that the images in SSDD already contain noise. The experimental parameters and environment are consistent with Section 4. In addition, the training parameters, including the epoch number and the initial parameter values, are set identically for the three SAR imaging networks.

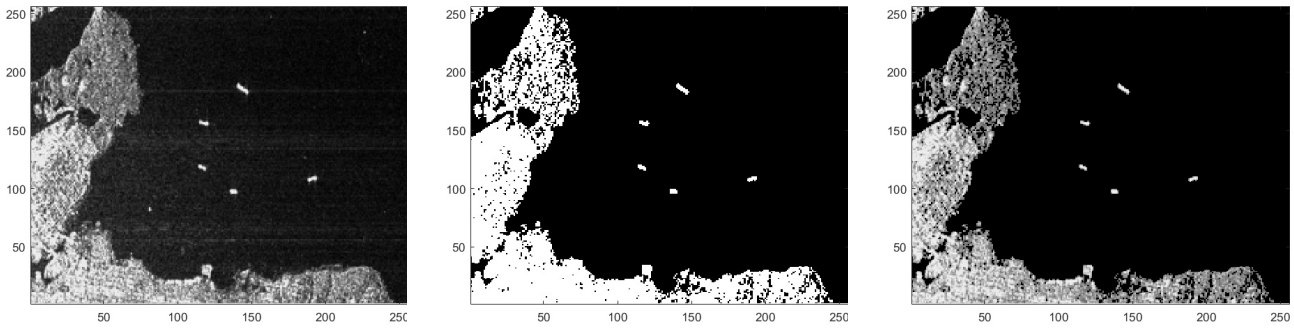


Figure 8. The process of the real scene label. (**a**) Original SAR image after cropping. (**b**) Mask obtained by the maximum connection domain algorithm. (**c**) Label of real scene.

Figures 9 and 10 display the recovered images of two different scenes indexed 724 and 751 in SSDD by CSA, L_1-De, CSA-Net, SR-CSA-Net, and the proposed method, respectively (the horizontal axis is the range direction). For full-sampled echoes, compared with CSA, the other four methods can productively suppress noise in the SAR images. However, both

L_1-De and CSA-Net also inadvertently suppress some target points and their scattering intensities. At a high down-sampling ratio with DSR = 72%, L_1-De and CSA-Net can restrain the energy dispersion and achieve acceptable recovered images. When DSR is decreased to 0.49, it is evident that L_1-De and CSA-Net are no longer competent for image recovery. SR-CSA-Net is capable of suppressing the ambiguity phenomenon while preserving the original scattering intensity of the real scene. But as the down-sampling ratio decreases, the contours of its recovered images become blurred. Furthermore, the recovered images of our complex-valued network have less noise and clearer edges than the other four methods.

The average values of the quantitative indicators of the recovered images are shown Tables 4 and 5. In the real scene experiments, CSA has the shortest computation time, which is only 3 ms, while L_1-De has the longest computation time. It is seen that compared with L_1-De, the three networks not only exhibit better imaging performance but also meet the real-time imaging requirement. Moreover, CSA-Net and L_1-De are seriously affected by down-sampling, with their quantitative indicators gradually approaching CSA as the DSR decreases. Table 4 indicates that the proposed method achieves approximately 2 dB and 6 dB gains over SR-CSA-Net and CSA-Net, respectively. In conclusion, our complex-valued network shows the optimal performance in the real scene reconstruction from both full- and down-sampled data.

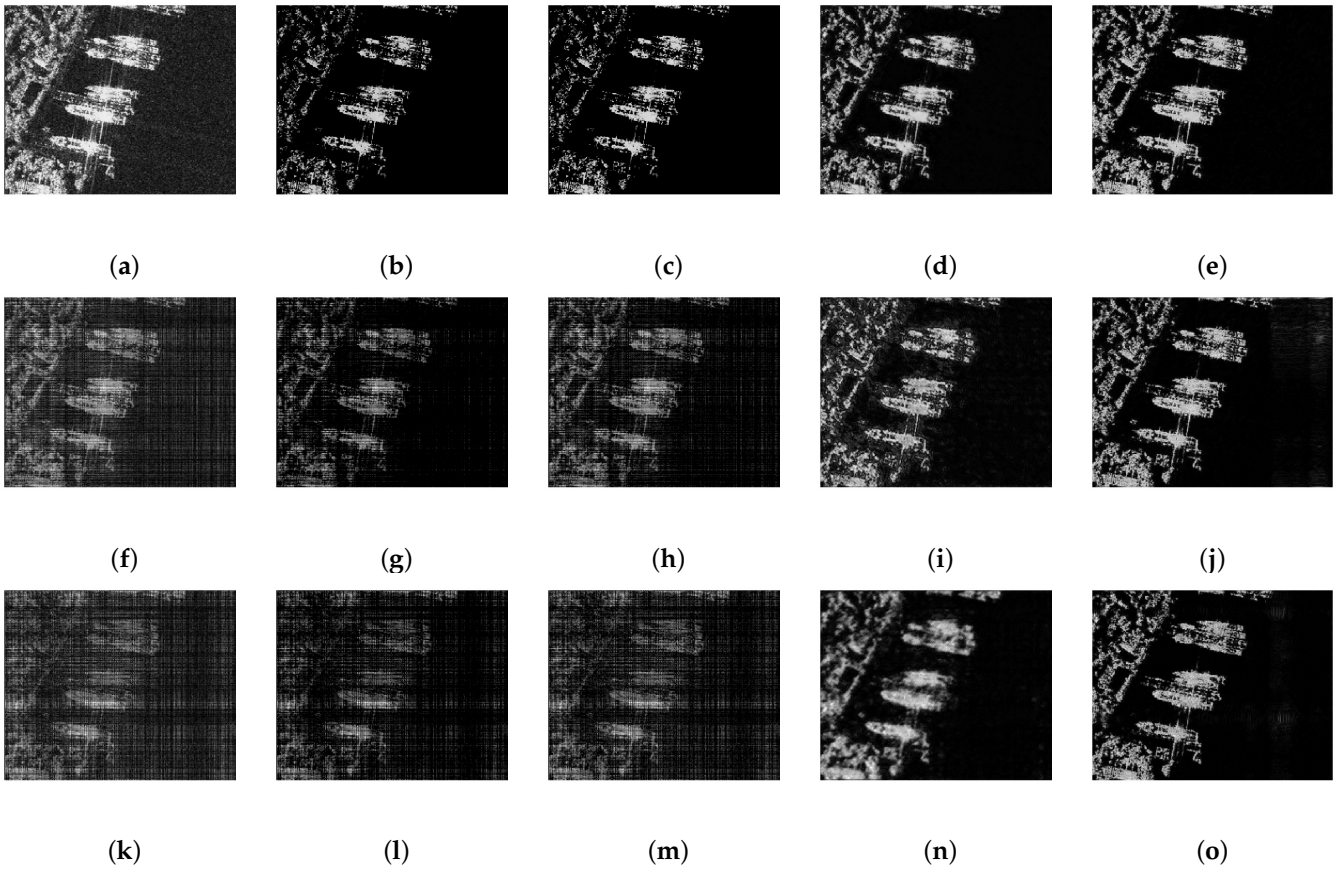


Figure 9. Recovered images of real scene 1 from down-sampled data by (**a**,**f**,**k**) CSA, (**b**,**g**,**l**) L_1-De, (**c**,**h**,**m**) CSA-Net, (**d**,**i**,**n**) SR-CSA-Net, and (**e**,**j**,**o**) our complex-valued network, respectively. From top to bottom column, the DSRs are 100%, 72%, and 49%, respectively.

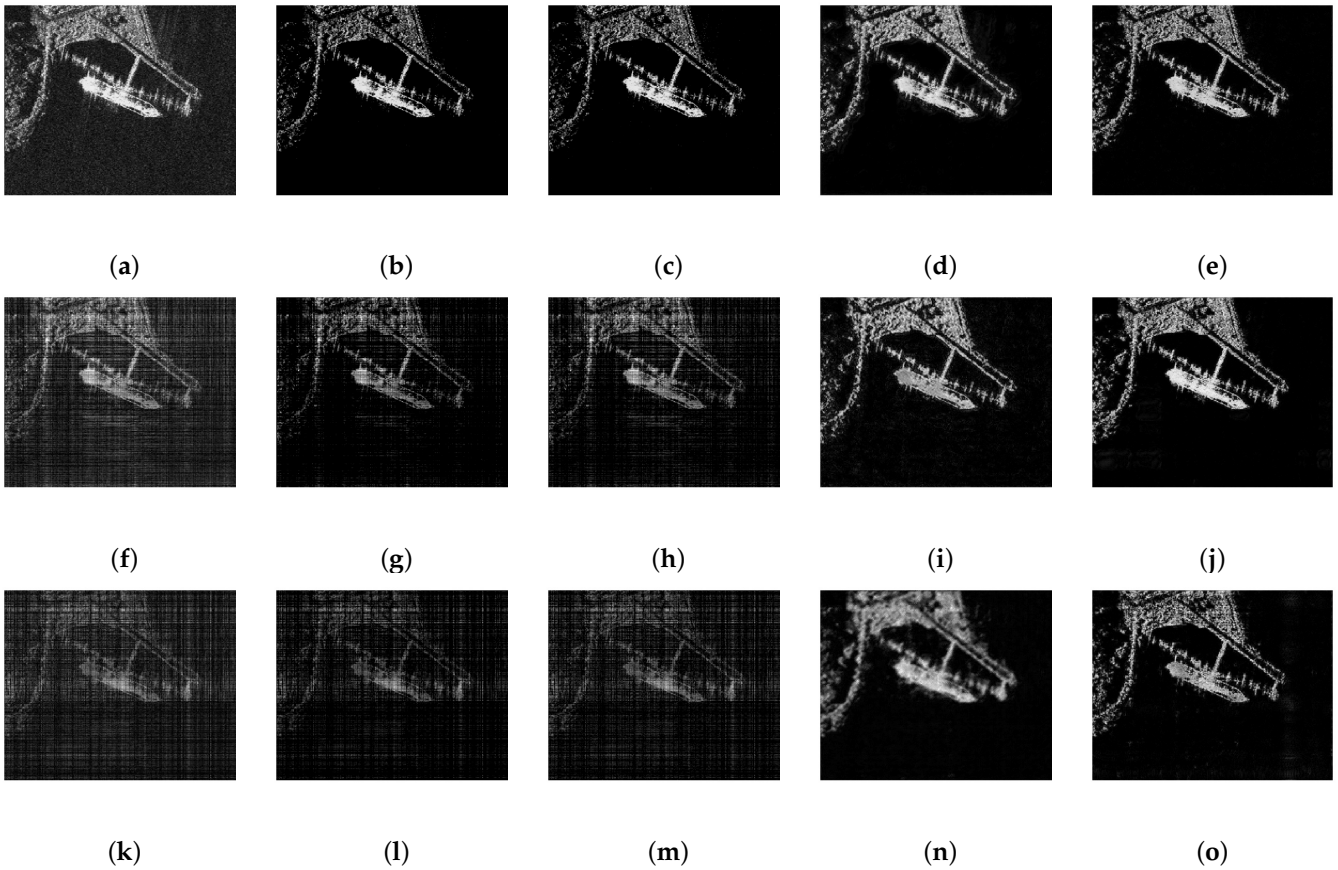


Figure 10. Recovered images of real scene 2 from down-sampled data by (**a**,**f**,**k**) CSA, (**b**,**g**,**l**) L_1-De, (**c**,**h**,**m**) CSA-Net, (**d**,**i**,**n**) SR-CSA-Net, and (**e**,**j**,**o**) our complex-valued network, respectively. From top to bottom column, the DSRs are 100%, 72%, and 49%, respectively.

Table 4. Performance comparison of real scene 1 with different down-sampling ratios.

Method	DSR = 100%		DSR = 72%		DSR = 49%		Time (ms)
	NMSE	PSNR (dB)	NMSE	PSNR (dB)	NMSE	PSNR (dB)	
CSA	0.5002	12.53	0.5367	12.22	0.6570	11.34	3.2
L_1-De	0.3228	14.43	0.5007	12.52	0.6435	11.43	149
CSA-Net	0.3091	14.62	0.4852	12.66	0.6179	11.61	9.8
SR-CSA-Net	0.1253	18.54	0.2177	16.14	0.2537	15.47	14.1
The proposed method	0.0792	20.53	0.1311	18.34	0.1663	17.31	21.4

Table 5. Performance comparison of real scene 2 with different down-sampling ratios.

Method	DSR = 100%		DSR = 72%		DSR = 49%		Time (ms)
	NMSE	PSNR (dB)	NMSE	PSNR (dB)	NMSE	PSNR (dB)	
CSA	0.6860	13.72	0.7111	13.56	0.7234	13.48	3.3
L_1-De	0.3819	16.26	0.5120	14.99	0.6625	13.87	150
CSA-Net	0.3673	16.43	0.5002	15.09	0.6543	13.92	9.8
SR-CSA-Net	0.1128	21.56	0.1638	19.94	0.2600	17.93	14.3
The proposed method	0.0829	22.89	0.1216	21.23	0.1728	19.71	21.6

6. Conclusions

We develop a new complex-valued CNN-based approximated observation sparse SAR imaging method. The typical approximated observation-based sparse imaging method is time-consuming due to multiple iterative operations, and it is challenging to obtain the optimal parameters. Therefore, we map the IST algorithm into a deep network form, i.e., single-path complex-valued CNN, to decrease the number of iterations. Extensive experiments show that our complex-valued network achieves accurate sparse reconstruction

of the considered scene from both full- and down-sampled echoes. Compared with CSA, L_1-De, CSA-Net, and SR-CSA-Net, our complex-valued network shows better performance in SAR imaging quality improvement, especially in the down-sampling cases. In addition, it further reduces the computational time for sparse imaging to a level comparable to that of the MF-based methods.

Since the proposed network for solving the L_1-norm regularization problem associated with sparse representation is very common and effective, a future research orientation is to build deep networks using other optimization algorithms, like CAMP [32] and BiIST [33]. Additionally, the proposed method only considers the side-look SAR imaging mode of static targets. In future work, we will further develop SAR imaging networks for moving targets under high-squint conditions.

Author Contributions: Conceptualization was done by H.B.; Z.J. and L.L. developed the methodology; validation was conducted by Z.J. and L.L.; Z.J. and L.L. prepared the original draft; H.B. performed the review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 62271248, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20230090, and in part by the Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of China under Grant KLSMNR-K202303.

Data Availability Statement: The real scene image used in this study is the SSDD dataset, which can be found at https://drive.google.com/file/d/1glNJUGotrbEyk43twwB9556AdngJsynZ/view?usp=sharing (accessed on 14 October 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this paper:

SAR	Synthetic Aperture Radar
ISAR	Inverse SAR
PolSAR	Polarimetric SAR
TomoSAR	SAR tomography
DL	Deep Learning
IST	Iterative Soft Thresholding
CNN	Convolutional Neural Network
PSRI-Net	Parametric Super-Resolution Imaging Network
CSA	Chirp-scaling algorithm
MF	Matched Filtering
CS	Compressive Sensing
FMCW	Frequency Modulation Continuous Wave
WDA	Wavenumber Domain Algorithm
SLC	Single-Look Complex
SR-IST-Net	Sparse Representation-based ISTA-Net
RDA	Range Doppler Algorithm
ADMM	Alternating Direction Method of Multipliers
BN	Batch Normalization
MC-ADMM	Multicomponent ADMM
SCR	Signal-to-Clutter Ratio
1-D	One-dimensional
RIP	Restricted Isometry Property
2-D	Two-dimensional
CAMP	Complex Approximated Message Passing
ReLU	Rectified Linear Unit
3-D	Three-dimensional
SNR	Signal-to-Noise Ratio

C-R	Cauchy–Riemann
DSR	Down-Sampling Ratio
NMSE	Normalized Mean Square Error
PSNR	Peak SNR
SSDD	Open SAR Ship Detection Dataset

References

1. Zhang, B.; Hong, W.; Wu, Y. Sparse microwave imaging: Principles and applications. *Sci. China Inf. Sci.* **2012**, *55*, 1722–1754. [CrossRef]
2. Donoho, D. Compressed sensing. *IEEE Trans. Inform. Theory* **2006**, *52*, 1289–1306. [CrossRef]
3. Wu, Y.; Hong, W.; Zhang, B.; Jiang, C.; Zhang, Z.; Zhao, Y. Current developments of sparse microwave imaging. *J. Radars* **2014**, *3*, 383–395. [CrossRef]
4. Baraniuk, R.; Steeghs, P. Compressive radar imaging. In Proceedings of the 2007 IEEE Radar Conference, Waltham, MA, USA, 17–20 April 2007; pp. 128–133.
5. Patel, V.; Easley, G.; Healy, D.; Chellappa, R. Compressed synthetic aperture radar. *IEEE J. Sel. Topics Signal Process.* **2010**, *4*, 244–254. [CrossRef]
6. Jiang, C.; Zhang, B.; Zhang, Z.; Hong, W.; Wu, Y. Experimental results and analysis of sparse microwave imaging from spaceborne radar raw data. *Sci. China Inf. Sci.* **2012**, *55*, 1801–1815. [CrossRef]
7. Fang, J.; Xu, Z.; Zhang, B.; Hong, W.; Wu, Y. Fast compressed sensing SAR imaging based on approximated observation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 352–363. [CrossRef]
8. Bi, H.; Wang, J.; Bi, G. Wavenumber domain algorithm-based FMCW SAR sparse imaging. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7466–7475. [CrossRef]
9. Bi, H.; Bi, G.; Zhang, B.; Hong, W.; Wu, Y. From theory to application: Real-time sparse SAR imaging. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2928–2936. [CrossRef]
10. Wang, Y.; Qian, K.; Zhu, X. Efficient SAR tomographic inversion via sparse bayesian learning. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium (IGASS), Brussels, Belgium, 12–16 July 2021; pp. 4830–4832.
11. Zhang, S.; Dong, G.; Kuang, G. Matrix completion for downward-looking 3-D SAR imaging with a random sparse linear array. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 1994–2006. [CrossRef]
12. Rao, W.; Li, G.; Wang, X. ISAR imaging via adaptive sparse recovery. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium (IGASS), Melbourne, VIC, Australia, 21–26 July 2013; pp. 121–124.
13. Chierchia, G.; Cozzolino, D.; Poggi, G.; Verdoliva, L. SAR image despeckling through convolutional neural networks. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGASS), Fort Worth, TX, USA, 23–28 July 2017; pp. 5438–5441.
14. Mousavi, A.; Baraniuk, R. Learning to invert: Signal recovery via deep convolutional networks. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017; pp. 2272–2276.
15. Gao, J.; Deng, B.; Qin, Y.; Wang, H.; Li, X. Enhanced radar imaging using a complex-valued convolutional neural network. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 35–39. [CrossRef]
16. Rittenbach, A.; Walters, J.P. RDAnet: A deep learning based approach for synthetic aperture radar image formation. *arXiv* **2021**, arXiv: 2001.08202.
17. Pu, W. Deep SAR imaging and motion compensation. *IEEE Trans. Image Process.* **2021**, *30*, 2232–2247. [CrossRef]
18. Meraoumia, I.; Dalsasso, E.; Denis, L.; Abergel, R.; Tupin, F. Multitemporal speckle reduction with self-supervised deep neural networks. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5201914. [CrossRef]
19. Lin, H.; Jin, K.; Yin, J.; Yang, J.; Zhang, T.; Xu, F.; Jin, Y. Residual in residual scaling networks for polarimetric SAR image despeckling. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5207717. [CrossRef]
20. Zhang, J.; Ghanem, B. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1828–1837.
21. Yonel, B.; Mason, E.; Yazıcı, B. Deep learning for passive synthetic aperture radar. *IEEE J. Sel. Topics Signal Process.* **2018**, *12*, 90–103. [CrossRef]
22. Zhang, H.; Ni, J.; Xiong, S.; Luo, Y.; Zhang, Q. SR-ISTA-Net: Sparse representation-based deep learning approach for SAR imaging. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 4513205. [CrossRef]
23. Kang, L.; Sun, T.; Luo, Y.; Ni, J.; Zhang, Q. SAR imaging based on deep unfolded network with approximated observation. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5228514. [CrossRef]
24. Wei, Y.; Li, Y.; Ding, Z.; Wang, Y.; Zeng, T.; Long, T. SAR parametric super-resolution image reconstruction methods based on ADMM and deep neural network. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10197–10212. [CrossRef]
25. Li, M.; Wu, J.; Huo, W.; Jiang, R.; Li, Z.; Yang, J.; Li, H. Target-oriented SAR imaging for SCR improvement via deep MF-ADMM-Net. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5223314. [CrossRef]

26. Zhang, H.; Ni, J.; Li, K.; Luo, Y.; Zhang, Q. Nonsparse SAR scene imaging network based on sparse representation and approximate observations. *IEEE Trans. Geosci. Remote Sens.* **2023**, *15*, 4126. [CrossRef]
27. Budillon, A.; Johnsy, A.C.; Schirinzi, G.; Vitale, S. SAR tomography based on deep learning. In Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium (IGASS), Yokohama, Japan, 28 July–2 August 2019; pp. 3625–3628.
28. Wang, Y.; Liu, C.; Zhu, R.; Liu, M.; Ding, Z.; Zeng, T. MAda-Net: Model-adaptive deep learning imaging for SAR tomography. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5202413. [CrossRef]
29. Li, H.; Xu, J.; Song, H.; Wang, Y. PIN: Sparse aperture ISAR imaging via self-supervised learning. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 3502905. [CrossRef]
30. Raney, R.; Runge, H.; Bamler, R.; Cumming, I.; Wong, F. Precision SAR processing using chirp scaling. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 786–799. [CrossRef]
31. Shi, W.; Ling, Q.; Yuan, K.; Wu, G.; Yin, W. On the linear convergence of the ADMM in decentralized consensus optimization. *IEEE Trans. Signal Process.* **2014**, *62*, 1750–1761. [CrossRef]
32. Bi, H.; Zhang, B.; Zhu, X.; Hong, W.; Sun, J.; Wu, Y. L_1 regularization-based SAR imaging and CFAR detection via complex approximated message passing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3426–3440. [CrossRef]
33. Bi, H.; Li, Y.; Zhu, D.; Bi, G.; Zhang, B.; Hong, W.; Wu, Y. An improved iterative thresholding algorithm for L_1-norm regularization based sparse SAR imaging. *Sci. China Inf. Sci.* **2020**, *63*, 219301. [CrossRef]
34. Trabelsi, C.; Bilaniuk, O.; Zhang, Y.; Serdyuk, D.; Subramanian, S.; Santos, J.; Mehri, S.; Rostamzadeh, N.; Bengio, Y.; Pal, C. Deep complex networks. *arXiv* **2018**, arXiv:1705.09792.
35. Li, J.; Qu, C.; Shao, J. Ship detection in SAR images based on an improved faster R-CNN. In Proceedings of the 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSARDATA), Beijing, China, 1–6 November 2017.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.